

基于生成词库理论的汉语同素词辨析*

曾立英

(中央民族大学 国际教育学院, 北京 100081)

[摘要] 针对汉语教学中同素词易混淆和大纲及教材中同素词词义计数式词库的问题, 文章提出基于生成词库理论来统一地、形式化地分析汉语同素词的词汇信息表征框架, 举例分析了汉语同素词的论元结构、事件结构、物性结构和词汇承继结构这四个层面, 同时为了抓住同素词语义的创造性, 分析了同素词的类型强制、成分共组和选择约束等生成设置, 以发掘同素词的生成机制。

[关键词] 同素词; 生成词库理论; 语义表征; 生成设置

中图分类号: H14 **文献标识码:** A **文章编号:** 1672-1306(2017)04-0047-09

一、引言

国际汉语教学中, 同素词教学是词汇教学的重点和难点。关于现代汉语同素词, 大致有比较严格和宽泛的两种定义, 比较严格的定义以丁勉哉^①为代表, 认为同素词是一种词素相同而结构次序相反的词, 即同素逆序词, 比如“语言”和“言语”等。而以常敬宇^②、刘叔新^③为代表的学者则对同素词定义较宽泛, 常敬宇认为, 在现代汉语词汇中, 有一种介于词和词族之间的词汇单位叫作同素词群。同素词群里的词彼此在词素上都存在着同一性的联系, 所以把词素相同的一组词, 彼此互称为同素词。如以“火”为词素构成的词有“炉火、煤火、焰火、烈火、怒火”等。刘叔新提出“同语素词语族”这个术语, 简称为“同素族”, 认为同素族含有同样一个语素, 即含有同一材料的共同意义成分, 而彼此在意义上相互因应, 使相应的词语单位形成一种聚合的组织, 比如“地、地面、地底、大地、落地、坠地、脚踏实地、平地一声雷”等, 称之为“同素族”。本文发现汉语教学中, 同素词之间语义的交织, 对外国学生容易形成困扰, 故采取比较宽泛的定义, 选取刘叔新的定义, 主要考察意义比较实在的共同语素的构词情况。

外国学生经常混淆同素词, 比如“生”与“出生”“经验”与“经历”等, 这两组词在 HSK 动态作文语料库^④中的偏误分别有 10 例、92 例, 括号里为正确词语:

1) 孩子生了以后每天学会各种各样的东西。(出生)

2) 我来华三年多了, 我在北京语言文化大学学汉语的过程中, 我经验了很多有意思的和不太高兴的事情。(经历)

学生经常出现如例 1) 中将“出生”误用为“生”的单向错误, 这是由于这两个词的论元结构(argument structure)不同所导致的, “生”和“出生”这两个动词所带的论元数目是不一样的, “生”可以带两个论元,

* 作者简介: 曾立英, 女, 湖北枝江人, 中央民族大学教授, 研究方向为汉语词汇学与词汇教学。

基金项目: 国家社会科学基金项目“现代汉语类词缀的定量与定性研究”(11BYY085); 国家自然科学基金项目“文本语言表达达到概念关系的映射方法研究与资源建设”(61375074)。

① 丁勉哉. 同素词的结构形式和意义的关系[J]. 学习月报, 1957, (2).

② 常敬宇. 同素词简论[J]. 世界汉语教学, 1985, (2).

③ 刘叔新. 汉语描写词汇学[M]. 北京: 商务印书馆, 1990: 359.

④ 网址: <http://202.112.195.192:8060/hsk/login.asp>. 查阅日期: 2016-6-10, 除非特别说明, 本文偏误例句均出于此语料库。

“出生”只能带一个论元。“经验”与“经历”都可以做动词和名词,学生经常出现当用“经历”却误用为“经验”的情况。杨寄洲、贾永芬^①只是将这种情况解释为:“动词‘经验’有‘经历’的意思,但是常用的是‘经历’;名词‘经历’有‘经验’的意思,但是常用的是‘经验’。”这样的解释有些重复。

目前对外汉语教学界关于同素词的研究主要聚焦于单双音节同义同素词,并且分词性进行了研究,如李泉^②、吴颖^③等对同素近义单双音节形容词的研究,程娟、许晓华^④、季瑾^⑤、张平^⑥等对同素单双音同义动词的研究,刘春梅^⑦、于洋^⑧等对同素单双音同义名词的偏误或混淆的研究。这些研究对汉语教学均有所帮助,但由于汉语的词缺乏词形变化,汉语教学中,学生无法立即从词形变化中分辨出词性,还有一些词是兼类词,之前的辨析方法大多是针对某种词性对同义同素词的辨析,是局部的解释,本文基于生成词库(Generative Lexicon)理论,寻求一套简洁的系统的语义表征和生成设置来探讨汉语同素词,给予同素词以统一的动态的解释,并且这种解释具有递归性。

二、针对汉语教材和大纲的同素词词库的构建

(一) 汉语同素词词库的构建与生成词库对照

同素词在教材、大纲中出现频率高,值得重视。我们选择了北京语言大学出版社的系列教材《成功之路》^⑨进行统计,发现生词表共 7367 个词语,由 2737 个语素构成,其中 108 个语素能构成同素词 20 条以上,如表 1 所示:

表 1 《成功之路》生词中构词频率高的共同语素

共同语素 构词数	共同语 素个数	共同语素枚举
80~82	1	一
70~79	2	人、大
60~69	3	不、生、心
50~59	1	意
40~49	13	力、气、天、体、来、有、出、自、成、上、学、然、面
30~39	24	手、好、地(dì)、行(xíng)、无、发、动、小、老、家、情、下、开、实、分、口、理、重、国、同、于、过、外、对
25~29	25	感、年、当、事、前、风、中、以、名、明、应、打、相、高、会、本、美、说、业、可、回、身、法、后、山
20~24	39	为、日、物、主、代、如、是、水、定、时、流、关、花、语、差、节、起、量、安、养、边、点、新、亲、多、车、进、观、神、提、性、制、现、样、信、工、包、白、想

① 杨寄洲、贾永芬.1700 对近义词语用法对比[M].北京:北京语言大学出版社,2005:712.
② 李泉.同义单双音形容词对比研究[J].世界汉语教学,2001,(4).
③ 吴颖.同素近义单双音节形容词的差异及认知模式[J].语言教学与研究,2009,(4).
④ 程娟,许晓华.HSK 单双音同义动词研究[J].世界汉语教学,2004,(4).
⑤ 季瑾.HSK 甲级单双音同义动词部分不可替换的类型探析[J].语言教学与研究,2005,(5).
⑥ 张平.同素同义单双音节动词的词义对应计算与分析[J].语言文字应用,2011,(3).
⑦ 刘春梅.留学生单双音同义名词偏误统计分析[J].语言教学与研究,2007,(3).
⑧ 于洋.CSL 学习者同素同义单双音名词混淆分布特征及其成因[J].语言教学与研究,2015,(6).
⑨ 邱军.成功之路[M].北京:北京语言大学出版社,2008.

我们还统计了《汉语水平词汇与汉字等级大纲》^①8822条词语，整理出同素名词258组、同素动词305组和同素形容词87组。该大纲中，构成同素名词频率最高的共同语素前10位是：“人(60)^②、地 di(40)、力(36)、面(34)、气(32)、事(30)、体(30)、年(28)、心(27)、方(26)”；构成同素动词频率最高的共同语素前11位是：“动(32)、发(28)、行(xíng)(25)、出(23)、解 jiě(20)、收(19)、交(18)、进(18)、开(18)、用(18)、转 zhuǎn(18)”；构成同素形容词频率最高的共同语素前13位是：“实(15)、平(13)、好(11)、密(10)、新(9)、正(9)、安(8)、静(8)、美(8)、难(8)、优(8)、严(8)、沉(8)”。现在表2中举例说明以该大纲为基础构成的同素词词库中部分共同语素的构词情况，并标出每个词在大纲中的级别，便于同素词的教学。

表2 《汉语水平词汇与汉字等级大纲》中的同素构词举例

共同语素	构词数	词 语	词性
部(乙)	18	部分(甲)、部门(乙)、部队(乙)、部长(乙)、部署(丙)、部件(丁)、部位(丁)、干部(甲)、全部(甲)、北部(乙)、东部(乙)、南部(乙)、内部(乙)、西部(乙)、局部(丙)、外部(丙)、中部(丙)、支部(丁)	同素名词
窗(甲)	4	窗户(甲)、窗口(丙)、窗帘(丙)、窗台(丙)	
想(甲)	28	想念(乙)、想象(乙)、猜想(丙)、幻想(丙)、回想(丙)、梦想(丙)、设想(丙)、构想(丁)、空想(丁)、联想(丁)、妄想(丁)、着想(丁)	同素动词
忘(甲)	2	忘记(乙)、忘却(丁)	
好(甲)	11	好看(甲)、好吃(甲)、好听(乙)、好玩儿(乙)、好些(乙)、好多(丁)、友好(甲)、良好(乙)、美好(乙)、正好(乙)、要好(丙)	同素形容词
难(甲)	8	难看(乙)、难过(乙)、难受(乙)、难得(丙)、难堪(丁)、难免(丁)、艰难(丙)、疑难(丁)	

表1和表2通过《成功之路》教材和《汉语水平词汇与汉字等级大纲》，根据构词频率的高低，抽取和构建了关于同素词的词库，这些词库也可加入共同语素和词的义项信息。但这种计数式词库貌似完备，也存在一些问题，比如“适合”与“合适”，我们可以把词的义项和每一个语素的意义都列举出来，但是还是不能解决学生将二者混淆的问题，混淆不单纯是意义的问题，还和论元结构等相关。Pustejovsky认为计算语言学 and 理论语言学家大都把词库作为一个词义的统计集，用句法特征、语素和语义信息进行标注，这属于一种词义计数式词库(Sense Enumeration Lexicons)。Pustejovsky批评词义计数式词库的框架是自然语言语义学贫乏的模式，这样的词库在下列语义描写中存在不足：(1)词的创造性用法的解释；(2)词义的渗透性(permeability)，即词义不是原子(atomic)的定义，而是与该词的其他意义有重叠和参照；(3)多种句法形式的表示，单个的词义可以有多种句法实现。这3种不足同时也是汉语同素词计数式词库面临的问题。^③

在同素词的辨析中，我们遇到了很多棘手的问题，辨析过程缺乏一种生成机制的理解，比如大多同素词的辨析都是按照词对这种平面的角度来进行分析，本文拟从生成词库的角度更新有关同素词的教学观念，不是着力于微观的同素近义词的辨析工作，而是从生成的角度发现同素词生成的语义模式。同素词词义是如何联系的？同素词的语义信息和句法信息如何关联？这对于词义计数式词库提出了挑

① 国家对外汉语教学领导小组办公室.汉语水平词汇与汉字等级大纲[M].北京:北京语言学院出版社,1992.

② 人(60)括号中数目为“人”的构词数,后文依次类推。

③ Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIT Press, 1995: 39.

战,于是生成词库提出其目标是能递归地将意义分配到表达式中,说明同义、反义、多义和转喻等现象。

(二)汉语同素词词义交织对于词义计数式词库的挑战

Pustejovsky 区分了对比(contrastive)歧义和补足(complementary)歧义。对比歧义指同音异义现象,指某个词偶然带了几个明显的不关联的意义,比如汉语表植物的“花”组成的“花、花朵、桃花”和表“用、耗费”义的“花钱、花费、花销”的“花”是同音异义词,是对比歧义,可以用词义的计数式来处理。补足歧义是同一个词的词义具有逻辑关联。Pustejovsky^①指出:“尽管计数式词库足以处理对比歧义,但是不能处理多义的本质。为了审慎处理补足的歧义,我们必须谈及词义和组构性(compositi-onality)的每一个假设。为了掌握组构性,我们必须丰富语义组构表达,在这个原则上做成的语义学模型的概要,叫生成词库。”

同素词共同语素的多个义项之间是有逻辑关联的,也是词义渗透性的一个重要体现。比如汉语“部”既可以指“部分;部位”,又可以指“单位”,同素多义造成了留学生的混淆,如:

3)先读提纲和结语,然后再逐步精读主要部门。(部分)

4)我希望在导游的部分工作。(部门)

例 3)、4)中“部分”与“部门”的混淆,可以用生成词库的物性结构(qualia structure)来分析。物性结构,定义物体的性质,由构成(constitutive)、形式(formal)、功用和施成(agentive)4 个角色构成。“部分”注重形式角色,是形式上的“整体”与“局部”的关系;而“部门”注重功用角色,它是一种社会机构,是功用上的“整体”与“局部”的关系。

生成词库采用一种生成词汇的方法转向到语义学,随之影响到语法,希望达到的目标是达成语言中意义的模式,通过这种模式能假定语境中意义潜在的有限的数目,同时限制存储在词库中意义的数目。

三、生成词库下汉语同素词的语义表征

Pustejovsky 认为生成词库的语义表征(semantic representations)详细说明词汇信息的组织,其目标是给语言提供一种形式化的表述,既有表现力,又能灵活地抓住词汇创造性和词义引申现象的生成本质。由此提出了一种新的分解(decomposit-ion)观点,词项(lexical items)被分解为最小的结构形式或块(templates),比分解成特征(features)集合适。简言之,生成词库至少包括 4 个层面的语义表征:论元结构、事件结构、物性结构和词汇承继结构(lexical inheritance structure)。这 4 种结构是相互作用的。

Pustejovsky^②用论元主体(args-body)模式(schema)说明论元结构、事件结构和物性结构的组构机制,其中论元结构抓住谓词中的参与者(participants),事件结构抓住作为事件或复杂事件的谓词,主体部分由物性结构和受事件结构利用的物性限制(constraints)构成。生成词库引进了一种丰富的词汇信息知识表征框架,不同于目前流行的静态词库。

前面例 1)已经提到了论元结构,论元结构指某个词项所包括的论元数目和论元类型,以及它们是怎样在句法上实现的。如留学生还经常将“合适”与“适合”“变动”与“改变”“着想”与“想”弄混:

5)我自认为贵社的职务很合适我。(适合)

6)我用一个道理来讲就是自然变化,如山会变动,水会减少。(改变)

7)我们应当着想一下这个问题。(想/考虑)

例 5)中“合适”与“适合”的偏误也可从论元结构来解释,“适合”后可直接带[+有生]论元,而“合适”不能直接带论元,如果需要带这样的论元,可用介词“对”等引出论元,如例 5)可改成“贵社的职务对我很

① Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995.

② Pustejovsky, James. Type Theory and Lexical Decomposition[J]. Journal of Cognitive Science, 2006, (49).

合适”。例 6) 中“变动”与“改变”语义相近,《现代汉语词典》也将“变动”解释为“**动**①变化;②改变”,但二者所带论元的类型不同,“变动”所带论元多指“社会现象”,而“改变”所带论元语义范畴较宽,指“事物”均可,所以例 6) 中“变动”应为“改变”。再如例 7) 中“着想”用错,主要在于“着想”这个动词不可直接带论元,如果需要带[+有生]论元,句法表现上需用“为、替”等介词将这一论元引出,如《成功之路》(进步篇)中的“为客户着想”用例。

有意思的是“交流”,可以说“交流经验/技术”等,但例 8) 却不能说:

8) 我们吃地地道道的中国菜,交流很多中国人。(日本学生作文)

当例 8) 中的“交流”,需要带[+有生]论元时,应该用“和/与”等介词引出人物,故例 8) 应改为“和很多中国人交流”。

论元结构是通过事件结构来体现的,事件结构不仅满足词项的基本的事件类型,同时也丰富词项的内部的、子事件(subeventual)结构。基本的事件类型包括状态(state)、过程(process)和转变(transition)3种事件结构^①。事件结构可以有很多子事件(subeventual)结构,比如事件 E_3 可以是一个复杂的事件结构,包括 E_1 、 E_2 两个子事件结构,事件的核心可以用 * 表示。^② 复合事件内子事件之间的时间顺序有三种:第一种是前一事件的发生和结束都在后一事件发生前,用符号 $<\infty$ 来表示;第二种是复合事件的多个子事件同时进行,用 $o\infty$ 表示;最后一种是后一事件的发生晚于前一事件的发生,但早于前一事件的完成,用符号 $[\text{?}]o\infty$ 表示(Pustejovsky^③、张秀松 2009^④)。

和英语等别的语言相比,汉语动结式是比较特殊的结构,很多外国学生回避使用该结构,如学生常常不会运用“v 开”这个结构,只会用其中的一个单音节动词:

9) 听到他这么说,第二个和尚也不愿意离这里了。(离开)

10) 在回中国的飞机上,我开了送我来的弟弟的女朋友给我的礼物。(打开)

“离开、打开、分开、张开、公开、敲开、踹开、化开、顶开”等“v 开”构成的同素词,可看成是一个复杂的事件结构,包括两个子事件:事件₁指过程,事件₂指状态,事件₁一般先于事件₂发生,比如“离开”的事件结构可用下图 1 表示:

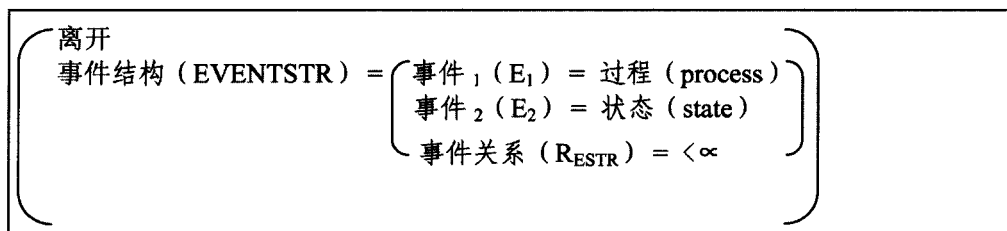


图 1 “离开”的事件结构

物性结构在前面例 3) 至例 4) 中已经提到了,例 2) 中“经验”应改为“经历”,“经验”侧重物性结构的功用(telic)角色,强调“正面的结果”,而且专指“由实践得来的知识或技能”,“经历”侧重于事件的过程,不关注结果的好坏。

这是生成词库一种重要的结构,和亚里士多德解释物体和关系的模式相关。物性结构反映了语义

① Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 61.

② Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 69.

③ Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 73.

④ 张秀松. 从生成词库论看汉语词的逻辑多义性[J]. 北方论丛, 2009, (3).

限制的概念,这种语义限制的概念隐藏在我们理解词的方式里。^① 其构成角色注重物体和构成成分之间的材料、重量、部分和组件元素之间的关系,形式角色区分物体在更大的“域(domain)”中的方向、量度、形状、维度、颜色和位置等,功用角色侧重物体的功能和目的,包括施事履行这个行为的目的、建造功用等,施成角色提示物体起源或者产生的因素,包括创造者、人工产品、自然类、使役链等。比如“窗口”和“窗户”,学生容易出现错误,如:

11)爱烟的人应该尊重不吸烟者的权力,在屋子里吸烟的话,把窗口打开等等办法来找共存。

(窗户)

例 11)中的“窗口”应为指实体的“窗户”,而不应该用为多指功用角色的“窗口”。“窗口”的物性结构偏重于其功用义,用在不同的领域有不同意义:用在服务业如“售票处、挂号室、银行”时,具有“营业”义;用在计算机操作系统,是“人机对话”的“窗口”;用到心灵、事物上,“窗口”具有比喻义。

“门口”与“窗口”在表层上看都属于“[房屋类]+口”构成的同素词,但是在物性结构上,“门口”着重指“门跟前”的处所义,而“窗口”如上所述则具有很多的引申义和比喻义。但是“门户”的功用角色比“门口”复杂,“门户”不仅可指形式角色的“门”,还可指“出入必经的要地”,可指“家”,可指“派别”,可指“门第”等,形成了补足的多义。从上面的分析可发现“门户、窗口”偏重物性结构的功用角色或施成角色,而“门口、窗户”则注重物性结构的形式角色或构成角色。

词汇承继结构识别某个词汇结构如何和类型格架(lattice)中其他结构关联,和它在整个词库组织中的贡献,同素词的很多聚合反映了承继性,比如“好”与“难”做形容词性语素时都能跟动词性语素如“看、吃、闻”等组成下列 A 组与 B 组词,具有一定的承继性:

A 组:难看、难吃、难闻、难听、难过、难办、难唱、难说、难受

B 组:好看、好吃、好闻、好听、好过、好办、好唱、好说、好受

再如下列 C 组与 D 组词中,“窗”构成的同素词与“门”构成的同素词也体现了承继性:

C 组:窗帘、窗框、同窗、窗口、窗户

D 组:门帘、门框、同门、门口、门户

在语义组构的操作中,论元、事件和物性类型必须符合被“类型系统”和词汇承继结构^②定义的合适性(well-formedness)条件。如果需要适应新奇的类型环境,词项会通过类型强制(type coercion)表达丰富的类型观。^③

四、汉语同素词的生成设置

Pustejovsky 指出,当某个词项与别的词项连接成短语或小句时,将通过生成设置(generative devices)产生扩展(extended)的意义。^④ 所谓“设置”,根据 Crystal(1980/2008),是生成语言学从数学借用的术语,指专门为分析提供的一种抽象设计。^⑤ 在简化语义的描写中,可以有几种设置,但生成词库的生成设置有自己的独特性,超出了词义计数的观念,指向了词义的结合。概括地说,生成词库是一种组织,产生扩展义的操作叫生成设置。这种生成操作又包括类型强制、成分共组(co-composition)和选择

① Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 86.

② 后来 Pustejovsky(2006: 39~76)从计算资源角度考虑,将词汇承继结构改为词汇类型结构(lexical typing structure)。

③ Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 59.

④ Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 61.

⑤ Crystal, David. A Dictionary of Linguistics and Phonetics (Sixth Edition) [M]. MA: Blackwell Publishing, 1980/2008: 141.

约束(selective binding)的语义转换(transformations)。生成设置这种观念支持语义模式化的解释观。

Pustejovsky 指出句法和语义范畴的同构(isomorphism)是不能在语言描写的所有层面维持的,即底层的语义类型和句法表征不是一一映射的,但还是存在典范的(canonical)句法表达式以让我们认识到语义类型。比如“相信”与“信任”这对同素词可从句法表现上予以区分,“信任”只能带名宾,其宾语语义限制在“人”或机构团体等,“相信”的宾语具有多种句法表现,可以是名宾,也可以是动宾和小句宾。

Pustejovsky 指出,词汇语义学面临的一个严肃问题是如何解释词项的系统歧义,他提出用类型强制和词汇管辖的类型提升(type-shifting)来说明这个问题。^①以 want 为例,want 的补足语至少有下列3种形式:

- 12)a. Mary wants John to leave. (S [+NF])
 b. Mary wants to leave. (VP [+INF])
 c. Mary wants a book. (NP)

Pustejovsky 认为例 12) 中 want 能出现在不同的句法环境中,与其说是 want 这个动词的意义发生了变化,还不如说是动词补足语的选择限制。Pustejovsky 将例 12) 的内部论元统一定型为命题(proposition),例 12)c 的 book 经历了类型强制,是 Mary wants a book to read,其中的 book 经历了类型提升。和 want 类似,enjoy、begin、finish 也产生了类似的语境意义,因此说这种动词的补足语的语义选择经历了类型提升操作,这种操作是通过动词的词汇管辖完成的,称之为类型强制。类型强制是一种语义装置,该装置将某个论元转变为某种功能所期待的类型,否则将导致错误。关于类型强制下的泛函运算,如公式 13) 所示:

13) 如果表达式 α 属于类型 c , 而且 β 属于类型 $\langle a, b \rangle$ 映射, 那么

- (I) 如果有类型 $c=a$, 那么 $\beta(\alpha)$ 属于类型 b 。
 (II) 如果有 $\sigma \in \Sigma_a$, 那么 $\sigma(\alpha)$ 组合能生成 a 型表达式, 那么 $\beta(\sigma(\alpha))$ 属于类型 b 。
 (III) 否则, 一个类型错误将产生。

类型强制是语义系统的一部分,允许我们抓住语义的创造性和词义引申现象,通过丰富的语义表征系统,类型强制才能准确地产生新义。

尽管前人已经提出每个表达式有一个类型阶(ladder),即语言中所有的表达式能被分配为一个基本类型。Pustejovsky 扩充了这一主张,认为每个表达式 α 可以有一系列的转变算子(shifting operators),称为 Σ_α ,它能改变表达式的类型和外延(denotation)。^②所谓类型提升,是指任何一个短语或名词,在语境中改变自己的类型。关于汉语的“赶论文、学钢琴”等结构的语义强制,Liu^③、宋作艳^④已经论述过。让我们来探讨一下汉语“想”所带的内部(internal)论元。汉语最近流行的一句话“我想静静”就有歧义,歧义的造成是“静静”既能以名词 N 的形式做“想”的宾语,表示“人物”;也能以 VP 的形式做“想”的宾语,表示“安静一下”。另外“想”还可以带小句宾语,如“我想他不会来”等,有意思的是,北京语言大学 BCC 语料库中还出现了“想 iphone”的用例^⑤,如:

① Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 110~111.

② Pustejovsky, James. Type Coercion and Lexical Selection[A]. In J. Pustejovsky (ed). Semantics and the Lexicon [C]. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1993.

③ Liu, Meichun. Lexical Information and Beyond: Constructional Inferences in Semantic Representation[A]. In: Proceedings of the 13th Pacific Asia Conference on Language, Information and Computation [C]. Taipei, 1999.

④ 宋作艳. 轻动词、事件与汉语中的宾语强迫[J]. 中国语文, 2011, (3).

⑤ 网址: <http://bcc.blcu.edu.cn/>.

14)我还加关注了,持续关注,想 iphone,我是想疯了,我要 iphone。(BCC 语料库)

例 14)中“想 iphone”中的 iphone 不再是一个单纯的名词,iphone 作为动词“想”的宾语,在“想”的“希望”意义控制下,其类型得到了提升,实际隐含了“想买/用 iphone”义,由此产生词义的创造性用法。

成分共组讨论动词逻辑的多义现象。简言之,共组描写一种结构在表层上可允许多种功能的运用,Pustejovsky 认为在成分共组中,某个短语的多重因素作为功能符(functions)在起作用,在组构中产生了词的新的非词汇化的意义。^①这也包括语义形式描写不足的例子通过方式(manner)共组、特征转写(transcription)和轻动词说明(light verb specification)等来丰富语境。共组允许我们利用功能词和短句中论元的语义信息,产生词库中没有列出来的创造性意义。

我们注意到“好”作为共同语素,能跟动词“办、唱、说、走、找”组成“好办、好唱、好说、好走、好找”,表示“容易”义,但是“好找”一词则有两种相反的意义,如:

15)春节过后讨论最热的话题就是“年后找什么工作”“哪些工作比较好找”等。(BCC 语料库)

16)娘娘,原来你在这里呀!让我一顿好找!(BCC 语料库)

例 15)中的“好找”是“容易找到”的意思,例 16)中的“好找”是“不容易找到”的意思。究其原因,在于例 16)中的“好找”有一个使役轻动词[CAUSE],和使役对象“我”共组产生后产生了语境中新的意义。

Pustejovsky^②针对形容词修饰语的多义现象提出了选择约束机制,Pustejovsky^③扩充了选择的概念,认为论元选择可以通过加强类型(types)系统和组构机制来说明。

汉语“快+名词性语素(Ng)”的组合也很多,如下列 E1 组和 E2 组。“慢+Ng”和“快+Ng”也有承继现象,如下列 E1 组和 F1 组都能成立;但“快+Ng”和“慢+Ng”也有不对称现象,如下列 E2 和 F2 组词,“慢+Ng”构词受到限制:

E1 组:快车、快班、快信、快件;

E2 组:快手、快嘴、快刀、快感、快意

F1 组:慢车、慢班、慢信、慢件

F2 组:*慢手、*慢嘴、*慢刀、*慢感、*慢意

E2 组能说,F2 组不能说,主要原因在于“快手、快刀、快感”中的“快”类似于例(22)的表达式,可看作是事件谓词,“快手”指“做事敏捷的人”,“快刀”指“切起来很快的刀”,“快感、快意”是“令人愉快/爽快”的感觉。而“慢手、慢刀、慢感”等因为偏向负面意义,故在该语境中,“慢”不能形成对中心语的合格的选择约束。

类型强制、成分共组和选择约束这三种语义转换很重要,因为它们抓住了典型的句法表达式之间语义的相关性。当经历语义组构的操作时,这三种语义转换能帮助我们抓住语义的创造性。

五、结 语

本文针对国际汉语教学中实际遇到的一些同素同义词的困惑,提出运用生成词库理论的论元结构、事件结构、物性结构和词汇承继结构等语义表征来分析“合适、适合”、“离、离开”、“部分、部门”等汉语同素词易混淆的问题。和传统的词义计数式词库对比,生成词库对汉语同素词的分解、组构,有着递归性特点,能将意义递归地分配到表达式中。

由于词在语境中经常产生创造义,于是本文分析了汉语同素词在词义结合时所体现的生成设置,指

① Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 61, 125~127.

② Pustejovsky, James. The Generative Lexicon[M]. Cambridge: MIP Press, 1995: 127~131.

③ Pustejovsky, James. Type Theory and Lexical Decomposition[J]. Journal of Cognitive Science, 2006, (49)

出了一些词语如“想 iPhone、好找、快手”等体现的类型强制、成分共组和选择约束等语义转换及其语义转变算子在转换中的作用。

本文运用生成词库理论对汉语同素词的语义表征和生成设置进行操作，在理论上和实际运用中对汉语同素词进行新的思考与分析，实现语义的深层概念结构与句法信息真正的对接，对于汉语教学和自然语言处理等均会有所帮助。

（这篇文章是笔者2015年在美国 University of California, Los Angeles 访学期间完成的，谨致谢忱。）

Discrimination of Chinese words with shared morphemes based on the Generative Lexicon Theory

ZENG Li-ying

(College of International Education, Minzu University of China, Beijing 100081, China)

Abstract: This paper aims to discriminate the confusing words with shared morphemes in Chinese Teaching and Learning, and proposes a unified and formative analysis of a knowledge representation framework for the lexical information of Chinese words with shared morphemes based on the Generative Lexicon Theory that avoid the inadequacies of sense-distinguished enumerative lexicons. The paper attempts to expound the semantic representations of Chinese words with shared morphemes at the levels of argument structure, event structure, qualia structure and lexical inheritance structure. A set of generative devices connects these four levels of representations, providing for the compositional interpretation of words in the context. Included in these generative operations are the type coercion, co-composition and selective binding. These semantic transformations are important to capture semantic creativity.

Key Words: words with shared morphemes; Generative Lexicon Theory; semantic representations; generative devices

[责任编辑:张黎玲]